Distributed Routing in Small Worlds - Evolution

Oskar Sandberg

Kleinberg's Result

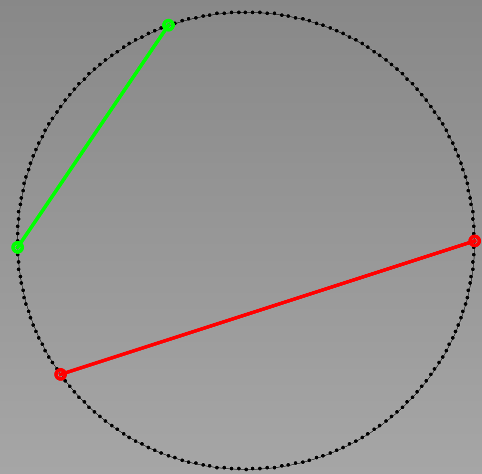
- We have a ring of nodes as a base graph.

Kleinberg's Result

- We have a ring of nodes as a base graph.
- We wish to add random edges to make it a small world, and we do so that the probability that node x chooses y for a neighbor is given by a kernel $\ell(x, y)$.

Kleinberg's Result

- We have a ring of nodes as a base graph.
- We wish to add random edges to make it a small world, and we do so that the probability that node x chooses y for a neighbor is given by a kernel $\ell(x, y)$.
- Kleinberg studied the class $\ell(x, y) = k|x - y|^{\alpha}$ for different α, with k as the normalizer.



Kleinberg's Result. cont.

$$\ell(x, y) = k|x - y|^{\alpha}$$

Kleinberg's Result. cont.

$$\ell(x, y) = k|x - y|^{\alpha}$$

- When $\alpha > -1$ there are too many long links, and we cannot find the right ones.

Kleinberg's Result. cont.

$$\ell(x, y) = k|x - y|^{\alpha}$$

- When $\alpha > -1$ there are too many long links, and we cannot find the right ones.
- When $\alpha < -1$ we can find long links taking us the right way, but not far enough.

Kleinberg's Result. cont.

$$\ell(x, y) = k|x - y|^{\alpha}$$

- When $\alpha > -1$ there are too many long links, and we cannot find the right ones.
- When $\alpha < -1$ we can find long links taking us the right way, but not far enough.
- **But** $\alpha = -1$ **is just right**!

Neighbor Selection

Kleinberg's result is somewhat surprising.

- Most network models do not allow for decentralized routing.

Neighbor Selection

Kleinberg's result is somewhat surprising.

- Most network models do not allow for decentralized routing.
- Yet is seems that social networks do have the property (Milgram).

Neighbor Selection

Kleinberg's result is somewhat surprising.

- Most network models do not allow for decentralized routing.
- Yet is seems that social networks do have the property (Milgram).
- Why should such networks arise in nature?

Neighbor Selection, cont.

- We derive an alternative criteria for efficient greedy routing.

Neighbor Selection, cont.

- We derive an alternative criteria for efficient greedy routing.
- Let $\ell(x, y)$ be such that it is proportional to the probability that greedy routing for y passes through x.

Neighbor Selection, cont.

- We derive an alternative criteria for efficient greedy routing.
- Let $\ell(x, y)$ be such that it is proportional to the probability that greedy routing for y passes through x.
- We can then achieve an upper bound on the expected routing time similar to Kleinberg's.

Neighbor Selection, cont.

Formally:

- Let $X_Z^Y(t)$ be a greedy walk through the network between random points Y and Z.

Neighbor Selection, cont.

Formally:

- Let $X_Z^Y(t)$ be a greedy walk through the network between random points Y and Z.
- If we fix $Z = z$, and let the final point be absorbing, we may view this as a Markov chain.

Neighbor Selection, cont.

Formally:

- Let $X_Z^Y(t)$ be a greedy walk through the network between random points Y and Z.
- If we fix $Z = z$, and let the final point be absorbing, we may view this as a Markov chain.
- Let $h(x, z) = \mathbf{P}(X_Z^Y(t) = x \text{ for some } t | Z = z)$ be the hitting probability of x when routing for z.

Neighbor Selection, cont.

Formally:

- Let $X_Z^Y(t)$ be a greedy walk through the network between random points Y and Z.
- If we fix $Z = z$, and let the final point be absorbing, we may view this as a Markov chain.
- Let $h(x, z) = \mathbf{P}(X_Z^Y(t) = x \text{ for some } t | Z = z)$ be the hitting probability of x when routing for z.
- $\tau = \sum_{\xi \neq z} h(\xi, z)$ is the expected length of a greedy route for z.

The Main Result

Theorem: *Let $N = 2^k$ with $k \geq 4$, and let $\ell(x, y)$ is such that:*

$$\ell(x, z) = \frac{h(|x - z|)}{\tau}.$$

A random graph constructed as above with these values has an expected greedy routing time between any two vertices $\leq 2k^2$.

We can prove this theorem when the base graph is a square toric grid or another transitive graph with similar properties.

Proof Sketch

- Let 0 be the destination, and number the other nodes $1, 2, \ldots, n-1$ after distance from destination.

Proof Sketch

- Let 0 be the destination, and number the other nodes $1, 2, \ldots, n-1$ after distance from destination.
- Let $k = \log_2 n$ and $F_1, F_2, \ldots, F_k$ be contiguous subsets such that we expected to spend as much time in each one.
  $F_1 = \{1, 2, \ldots, r_1\}$, $F_2 = r_1 + 1, r_1 + 2, \ldots, r_2$, etc, and $h(F_i) \approx h(F_j)$ for all i, j.

Proof Sketch

- Let 0 be the destination, and number the other nodes $1, 2, \ldots, n-1$ after distance from destination.
- Let $k = \log_2 n$ and $F_1, F_2, \ldots, F_k$ be contiguous subsets such that we expected to spend as much time in each one.
  $F_1 = \{1, 2, \ldots, r_1\}$, $F_2 = r_1 + 1, r_1 + 2, \ldots, r_2$, etc, and $h(F_i) \approx h(F_j)$ for all i, j.
- $\ell(F_i) \approx 1/k$.

Proof Sketch, cont.

Consider a phase F_k and assume it covers less then half the distance to the destination: $r_{k-1} > n/2$.

Proof Sketch, cont.

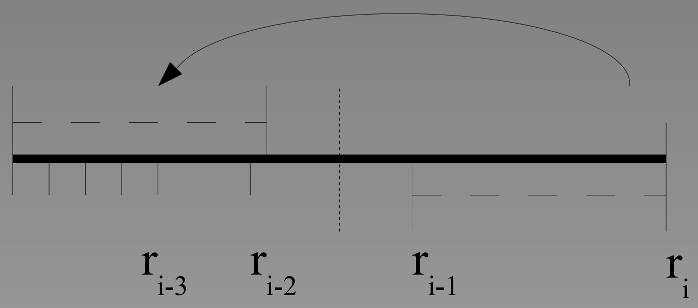
Consider a phase F_k and assume it covers less then half the distance to the destination: $r_{k-1} > n/2$.

- Then $r_k - F_k \subset \{0, 1, \ldots, r_{k-1}\}$:

Proof Sketch, cont.

Consider a phase F_k and assume it covers less then half the distance to the destination: $r_{k-1} > n/2$.

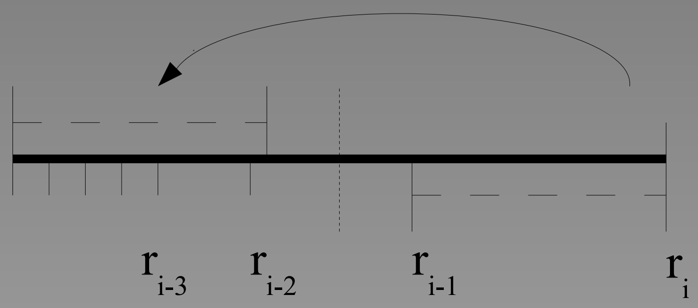
- Then $r_k - F_k \subset \{0, 1, \ldots, r_{k-1}\}$:



Proof Sketch, cont.

Consider a phase F_k and assume it covers less then half the distance to the destination: $r_{k-1} > n/2$.

- Then $r_k - F_k \subset \{0, 1, \ldots, r_{k-1}\}$:



- The probability of leaving in each step in F_k is at least $\ell(F_k) \approx 1/k$.

Proof Sketch, cont.

- If the probability of leaving F_k in each step is $1/k$, then $h(F_k) \leq k$.

Proof Sketch, cont.

- If the probability of leaving F_k in each step is $1/k$, then $h(F_k) \leq k$.
- It follows by construction $h(F_i) \leq k$ for all i, so $\tau \leq k^2$.

Proof Sketch, cont.

- If the probability of leaving F_k in each step is $1/k$, then $h(F_k) \leq k$.
- It follows by construction $h(F_i) \leq k$ for all i, so $\tau \leq k^2$.
- Thus we need only look at the case $F_{k-1} \leq n/2$.

Proof Sketch, cont.

- If the probability of leaving F_k in each step is $1/k$, then $h(F_k) \leq k$.
- It follows by construction $h(F_i) \leq k$ for all i, so $\tau \leq k^2$.
- Thus we need only look at the case $F_{k-1} \leq n/2$.
- By the exact same reasoning for the next phase, we need only consider $F_{k-2} \leq n/2^2$ and so on.

Proof Sketch, cont.

- In particular, we can eliminate all cases but $r_1 = n/2^{k-1} = 2$.

Proof Sketch, cont.

- In particular, we can eliminate all cases but $r_1 = n/2^{k-1} = 2$.
- But then $2 = |F_1| \geq h(F_1)$ (since $h(x) \leq 1$), whence $h(F_1) \leq k$ for all $k \geq 2$, and the result follows.

Proof Sketch, cont.

- In particular, we can eliminate all cases but $r_1 = n/2^{k-1} = 2$.
- But then $2 = |F_1| \geq h(F_1)$ (since $h(x) \leq 1$), whence $h(F_1) \leq k$ for all $k \geq 2$, and the result follows.
- The 2 in the statement of the theorem deals with the $\approx$ equality used in the begining.

Rewiring algorithm

Is this criteria more natural than Kleinberg's?

Rewiring algorithm

Is this criteria more natural than Kleinberg's?

Algorithm:

- Pick Y and Z uniformly at random from the set of vertices, and perform a greedy walk from Y to Z.

Rewiring algorithm

Is this criteria more natural than Kleinberg's?

Algorithm:

- Pick Y and Z uniformly at random from the set of vertices, and perform a greedy walk from Y to Z.
- For each vertex x in the walk, independently with probability $p < 1$ *re-wire* x's shortcut to point at Z.

Rewiring algorithm, cont.

- This algorithm defines a Markov chain on the set of shortcut configurations.

Rewiring algorithm, cont.

- This algorithm defines a Markov chain on the set of shortcut configurations.
- Under its stationary distribution, the shortcut at x is a sample of the destinations of queries that hit x.

Rewiring algorithm, cont.

- This algorithm defines a Markov chain on the set of shortcut configurations.
- Under its stationary distribution, the shortcut at x is a sample of the destinations of queries that hit x.
- Meaning:
  $\ell(x, z) = \mathbf{P}(Z = z | X_Z^Y(t) = x \text{ for some } t)$.

Rewiring algorithm, cont.

- This algorithm defines a Markov chain on the set of shortcut configurations.
- Under its stationary distribution, the shortcut at x is a sample of the destinations of queries that hit x.
- Meaning: $\ell(x, z) = \mathbf{P}(Z = z | X_Z^Y(t) = x \text{ for some } t)$.
- Bayes theorem, and Z being uniform then gives:

$$\ell(x, z) = \frac{\mathbf{P}(X_Z^Y(t) = x \text{ for some } t | Z = z)}{\tau}$$

Rewiring algorithm, cont.

- The re-wiring algorithm leads to graph with marginal shortcut distributions that have the property we seek.

Rewiring algorithm, cont.

- The re-wiring algorithm leads to graph with marginal shortcut distributions that have the property we seek.
- However, there are dependencies between the shortcuts at different vertices.

Rewiring algorithm, cont.

- The re-wiring algorithm leads to graph with marginal shortcut distributions that have the property we seek.
- However, there are dependencies between the shortcuts at different vertices.
- Proving that re-wiring algorithm leads to a routing friendly graph is an open problem, very strongly supported by experimental data.